

**UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF TEXAS
DALLAS DIVISION**

UNITED STATES OF AMERICA

§
§
§
§
§

v.

PRICE et al.

**No. 3:14-CR-00293-M-1
Judge Lynn**

DEFENDANTS' REPORT REGARDING THE STATE OF THE DISCOVERY

TO THE HONORABLE UNITED STATES DISTRICT JUDGE:

Counsel file this Notice of Discovery Issues to update the Court regarding the state of the Discovery and its accessibility to the defendants.

PROCEDURAL AND FACTUAL BACKGROUND OF THE DISCOVERY PRODUCTIONS AND DATABASE

During the course of the government's investigation, most likely commencing prior to the execution of the search warrants in June of 2011, the Government issued a number of grand jury subpoenas.¹ Presumably, the targets of those subpoenas produced whatever records existed and were within the scope of the grand jury subpoena.² The records and documents produced by the various targets of the subpoenas are "Native Files."

Upon receipt of the records, the government assigned a Bates range to each set of documents and subsequently labeled each page with a Bates number. Both the government and the defense agree that the Bates numbering system is instrumental to the task of organizing the material and understanding what material was produced, how it was produced, and by what entity. Without the Bates numbers, it is extremely difficult to identify and trace the source of the document. The government added Bates numbers on the face of the native files and produced the Bates labeled files to the Discovery Coordinating Attorney. These Bates labeled files are known as "Load Files." As ordered to do so by the Court, the government produced the load files in rolling productions. Accordingly, the discovery coordinator began uploading the load files to Summation to make them accessible to the defendants.

During the uploading process, multiple issues arose with the usability of the government's load files. It is counsels' understanding that these issues relate to coding of certain fields within the load files. These problems occurred in multiple productions and continued over a span of months. Eventually, in mid-January of 2016, the Discovery Coordinator was forced to abandon the use of the government-created load files and change course.

¹ The exact number of subpoenas is unknown to the defendants but has been estimated to be roughly 983 based on the prosecution's original discovery chart.

² To date, the prosecution has declined to make the subpoenas available to defense counsel so the scope of each subpoena is unknown to the defense.

The IT staff at the Public Defender's Office uploaded the native files as they were produced by the targets of the various subpoenas instead of the government-produced load files and replaced the improperly coded load files on Summation with the native version of each file. Replacing the load files caused delays in the process of making the full discovery available to the Defendants.

The Federal Public Defender's Office is working on a solution to the missing Bates number problem. It is counsels' understanding that that involves the recoding of fields and reprocessing of the load files. As of the date of this filing, it is unknown how long the recoding will take. To date, the complete discovery is not yet fully loaded on Summation.

SOURCING THE MATERIAL AND THE BATES NUMBERING ISSUES

As discussed above, the government's Bates numbering system is instrumental to organizing and sourcing the discovery. The defense, however, has not been able to use the government's Bates system because load files produced by the government were not loadable into Summation. When the load files were replaced by the native files, the government-assigned Bates numbers were eliminated from an overwhelming majority of the files on Summation. Thus, the defense has only had access to the native files that do not contain the government's Bates numbers.

The elimination of the Bates numbers from the files on Summation presented an entirely new problem. As such, the Discovery Coordinator created and implemented a plan to enable defense counsel to identify, as best as possible, the source of the material. Mr. Welaj, the Computer Systems Administrator at the FPD's Office, spent countless hours creating a master index for Government Discovery Production #s 5_1, 5_2, 5_3, 5_6, 7, and 19. To accomplish this task, Mr. Welaj had to manually correlate the ObjectName or DocID assigned by Summation with the government-assigned Bates numbers for 2,809,658 files referenced in the government's indices.³ The solution was successful to a certain degree. Unfortunately, Mr. Welaj's tireless work did not completely solve the problems associated with the use of the native files instead of the government's load files.

For example, Mr. Welaj was unable to create an additional master index for Government Discovery Production #4 – which consists of Materials Seized by Search Warrant – due to the sheer volume of the production set. If printed, the Government's Production 4 *Index* would be 2,158 pages, indexing 1,367,059 files. Further, the indices Mr. Welaj created are only helpful if the files are being viewed exclusively on Summation. Once the files are exported from Summation, as they would need to be in order to create Defense Exhibits, the files no longer contain any identifying number on the face of the document.

Additionally, Mr. Welaj's indices only allow the defense to determine what *entity* produced the files. Often times, however, multiple files were condensed and then uploaded to Summation as one file. In these instances, the beginning and end of each individual file cannot be determined and the Bates label for the interior pages – which were each individually Bates numbered by the government – cannot be identified. Thus, it is difficult to tell how many pages make up a specific document and what documents in the mass file are relevant to one another. As such, knowing which pages to extract in order to prepare an exhibit is extremely challenging.

³ The ObjectName and DocID are part of the unique identifying system used by Summation to identify each file.

For instance, Government Discovery Production 19 and 19A consist of “14 Banker’s Boxes of documents obtained by a ‘citizen’ who bid on a storage unit in default.”⁴ The 14 banker’s boxes held multiple documents that were separate and distinct from all the other documents in the set of banker’s boxes. The 14 banker’s boxes held, collectively, thousands of pages. The documents were scanned, en masse, and Bates labeled by the government. The load files, containing the government’s Bates numbers, were subsequently produced to the Discovery Coordinator and uploaded to Summation. The entirety of documents in the 14 banker’s boxes were, however, reduced to 37 *electronic* files. Most of those 37 files contain hundreds of pages that do not show the government-assigned Bates labels. Moreover, many of the files did not properly scan and appear to be nothing more than a gray X.

Further, Productions 5_1, 5_2, and 5_3 contain, collectively, 2,781,453 files.⁵ The majority of the 2,781,453 files are multiple pages in length, with many of the files consisting of hundreds of individual pages. Each page of the *load* files in Productions 5_1, 5_2, and 5_3 shows a government-assigned Bates label. The native files that the defense is viewing on Summation, however, do not contain those Bates numbers.

PROBLEMS WITH THE GOVERNMENT’S INDICES

In the government’s responses⁶ to the defendant’s discovery motions, the prosecution suggests that the government’s Bates numbering system, and the government-produced indices, contain sufficient information to enable the defendants to locate and source the entirety of the discovery materials. This is unfortunately not correct.⁷

For example, in the Government’s Response to Defendant Nealy’s Discovery Notice, the prosecution writes, “Exhibit 1 identifies each search location by the range of documents found from that location. For instance, documents marked with Bates labels in the range GE 0520 – 0712 were found at Dallas County’s offices.”⁸ The government has, however, produced two indices pertaining to the locations from which the GE-labeled files were seized. Those two indices are not consistent regarding the original locations of the GE files.

Specifically, on November 24, 2014, the government notified counsel by letter that Government’s Production #4 was available. Government’s Production 4 included an Index (***Production 4 Index***). On April 27, 2016, the Government produced an index titled Government Production 4 - GE Location Index⁹ (***GE Location Index***). **The ranges of the GE- numbered files attributed to each location in the *GE Location Index* do not coincide with the information listed in the *Production 4 Index*.** Further, defense counsel cannot determine, using only the information the government has made available, which of the two conflicting indices is accurate and which is not.

⁴ Government’s Production Letter 19 was sent to Counsel on September 16, 2015 and 19A was sent on January 28, 2016.

⁵ Production 5_1 contains 602,860 files; Production 5_2 contains 225,513 files and; Production 5_3 contains 1,953,080 native files.

⁶ See Government’s Response to Defendants’ First Joint Motion to Compel Discovery (dkt. 250) and Government’s Response to Defendant Nealy’s Discovery Notice (dkt. 248).

⁷ There have been lengthy discussions between the parties in the recent past and the government’s current position regarding the Bates numbering may not be the same as it was when the responses were filed on April 28, 2016.

⁸ Government’s Response to Defendant Nealy’s Discovery Notice, dkt. 248 @2.

⁹ The GE Location Index is Exhibit 1 (dkt. 249-1) of the Government’s Response to Defendant Nealy’s Discovery Notice (dkt. 248).

For example, the *GE Location Index* shows that the files in the GE 0520 - 0712 range were seized from Location A, Dallas County Offices. The *Production 4 Index*, however, shows that the GE-0704 - 0712 files were seized from Location B, the Residence of John Wiley Price. In the below examples, GE-0712 is used for purposes of illustration. The *GE Location Index*, as it pertains GE 0712, shows the following:

GE Range	Alpha	Location	Address
GE 0520 – 0712	A	Dallas County Offices	411 Elm Street, Dallas, Texas

Determining the source-location of the GEs using the *Production 4 Index* requires additional steps. First, the GE numbered file must be located on the GE section of the *Production 4 Index*. Next, it is necessary to determine from which set of Bates numbered files, and which specific “item” number, the GE file in question was produced. This information is listed in the last column of the GE section of the *Production 4 Index*.

For example, as it pertains to GE-0712, the *Production 4 Index* shows the following:

BEGINNING BATES NUMBER	ENDING BATES NUMBER (page count)	
GE-0712-0001	GE-0712-0221 (221)	1B-054, I-10

Specifically, in the above cited example, the *Production 4 Index* shows that GE-0712 is a 221-page document, extracted from the set of files Bates labeled with “1B-054-ITEM 10.” Lastly, in order to identify the location from where the “1B-054, ITEM 10” files were seized, it is necessary to locate the “1B-054, ITEM 10” Bates number in the body of the *Production 4 Index*.¹⁰

The body of the *Production 4 Index* shows the following files for “1B-054, ITEM 10”:

BEGINNING BATES NUMBER	ENDING BATES NUMBER (page count)
1B054-MISC ITEMS LOC B-ITEM 10-BLK FOLDER-EMPTY-000001	1B054-MISC ITEMS LOC B-ITEM 10-BLK FOLDER-EMPTY-000001 (1)
1B054-MISC ITEMS LOC B-ITEM 10-ELECTION FLYER-000001	1B054-MISC ITEMS LOC B-ITEM 10-ELECTION FLYER-000001 (1)
1B054-MISC ITEMS LOC B-ITEM 10-ELECTION FLYER-000002	1B054-MISC ITEMS LOC B-ITEM 10-ELECTION FLYER-000002 (1)
1B054-MISC ITEMS LOC B-ITEM 10-RED FOLDER W INVOICES-000001	1B054-MISC ITEMS LOC B-ITEM 10-RED FOLDER W INVOICES-000001 (1)
1B054-MISC ITEMS LOC B-ITEM 10-RED FOLDER W INVOICES-000002	1B054-MISC ITEMS LOC B-ITEM 10-RED FOLDER W INVOICES-000002 (1)

Thus, in the above cited example, the *Production 4 index* shows that the files labeled with “1B-054, ITEM 10” were seized from *Location B*, the Residence of John Wiley Price, and not Location A, Dallas County Offices, as shown in the *GE Location Index*.

Further, the *Production 4 Index* only shows a total of five files labeled with “1B-054, ITEM 10.” Each of those five files is shown to be one page in length and includes: an empty black folder, two election flyers, and two red folders with invoices. However, GE-0721 which purports to derive from 1B-054 ITEM 10 contains 221 pages of hand-written messages related to telephone calls. Clearly, it is physically impossible for a 221-page exhibit to have been extracted from five, one page files. This issue is not exclusive to GE-0712. To the contrary, there are multiple ranges of GE files that show one source location on the *GE Location Index* and a different source location on the *Production 4 Index*.

¹⁰ The body of the Production 4 Index identifies each individual file seized and gives a short description of the item.

There are other significant inconsistencies between the *GE Location Index* and the *Production 4 Index*. For example, the *GE Location Index* shows that GEs 4500 – 4550 came from five different locations – Locations H, I, K, M and N. Those five locations, however, are not ever referenced in the *Production 4 Index*. Further, the *Production 4 Index* shows that twelve different GE-labeled files came from a set documents with a Bates prefix that is not referenced in the body of the *Production 4 Index*.

For example, the *Production 4 Index* shows the following for GE-0697 and GE-0698:

BEGINNING BATES NUMBER	ENDING BATES NUMBER (page count)	
GE-0697-0001	GE-0697-0004 (4)	1B-041, 1-64
GE-0698-0001	GE-0698-0021 (21)	1B-031, 1-92

Bates numbers “1B-041” and “1B-031” are not, however, listed anywhere else in the *Production 4 Index* (or any other government-produced index). Thus, counsel does not have any way to identify the source of GE-0697 or GE-0698. There are other GE-labeled files that identify Bates prefixes not referenced elsewhere in the *Production 4 Index*. The source of those GE files is entirely unknown.

According to the Government’s letter to Counsel on November 25, 2014, “materials stamped with a ‘1B’ bates number are those seized pursuant to search warrants, and items stamped with a ‘GE’ came from the 1B materials and are likely to be used as evidence in the Government’s case-in-chief.” To the best of counsels’ knowledge, the documents labeled with a GE number were *only* produced by the government as GE labeled files. And, as illustrated in the GE-0712 example, *supra*, the page count and content of the GE files are often inconsistent with the page count and general description of the items in the body of the *Production 4 Index*. Thus, Counsel cannot determine how the GEs were created. Additionally, because the GE files were separated from the 1B items, Counsel cannot determine which of the 1B materials were combined to create a GE file by cross-referencing each GE file with a complete, uncultured set of documents seized from each location. Not having a complete set of the 1B (seized) material significantly undermines counsels’ ability to comprehend the source and context of the GE files.

Further, according to the *GE Location Index*, there are unexplained gaps in the GE numbering system.¹¹ For example, GE 0713 – GE 1019 are not listed. If files labeled with GE-0713 through GE-1019 exist, it is unlikely that the file has been produced in any government production because it appears that the GE files have only been produced as GEs.¹²

¹¹ GE files are not listed on the government’s indices for the following ranges: 0713 – 1019; 1100 – 1519; 1819 -2019; 2075 – 2520; 2581 – 3020; 3179 – 3519; 3599 – 3999.

¹² The gaps in the GE numbering system may be intentional so as to allow the government to supplement the GEs at a later date but, at the time this notice is filed, the gaps are unexplained.

EXAMPLES CITED IN THE GOVERNMENT'S RESPONSES TO THE DEFENDANTS' DISCOVERY FILINGS

The Government's Response to Defendant Nealy's Discovery Notice suggests that the Discovery is and has been searchable and manageable for the defendants.¹³ In support of that assertion, the government cites specific portions of the discovery material. For example, the government writes:

"For instance, if Nealy was interested in reviewing Business H's subpoena production, the index identifies the following Bates labels: SBP-1A-168; SBP-1A-486; SBP-1A-562; and Main-1A-235. Review of the materials within those productions would allow Nealy to identify internal, Business H emails discussing pursuit of the FTZ status. Put another way, Nealy's analogy to "10 banker boxes" (Notice at 2) ignores the government's Bates labeled production."¹⁴

Counsels' review of the above listed Bates-numbered documents yielded the following:

1. None of the files on Summation, pertaining to Business H, have government Bates numbers on the face of the documents.
2. One of those Bates labeled productions - Main-1A-235 - is missing entirely from Summation.
3. The remaining three Bates labeled productions - SBP-1A-168, SBP-1A-486 and SBP-1A-562 - contain, collectively, 49,572 files.
4. Of those 49,572 files, 47,342 are **not searchable** because the files were loaded as images after the attempt to load the government-created load files was abandoned.
5. In order to review the "internal, Business H emails discussing pursuit of the FTZ status", one would be required to click on all 47,342 *images* to determine which images contain emails relevant to the pursuit of FTZ status. Thus, the files pertaining to Business H, as they're available to the defendants, are the electronic equivalent of 10 banker's boxes of documents. To sift through 47,342 unsearchable images in search of emails regarding Business H's pursuit of FTZ status, with the limited resources and limited man power available to the defendants, would most likely take months to accomplish.

In the responses to the defendant's discovery related filings, the prosecution also cites examples that are slightly misleading or entirely inaccurate. For example, the prosecution writes:

"Just as with the records of Business H, discussed above, Nealy could examine the bank records for City Bank Account ending in 6372 by reviewing the index (Exh. 2) and identifying records marked with Bates labels SBP-1A-052, SBP-1A-088, SBP-1A-089, SBP-1A-090, and SBP-1A091. Counsel could limit her computer searches to those records or simply review them individually."¹⁵

The review of these documents is not as simple as the government's response suggests. Specifically, the government's response states "Nealy could examine the bank records for City Bank Account ending in 6372 by reviewing the index (Exh. 2) and identifying records marked with Bates labels SBP-1A-052, SBP-1A-088, SBP-1A-089, SBP-1A-090, and SBP-1A091."¹⁶ However, ironically, at the time the government's response was filed, Bates labels SBP-1A-088, SBP-1A-089, SBP-1A-090, and SBP-1A091 were not referenced anywhere in the government's indices, *including* Exhibit 2.¹⁷

¹³ As noted in footnote 6, the government's current position regarding the Bates numbering may not be the same as it was when the responses were filed on April 28, 2016.

¹⁴ Government's Response to Defendant Nealy's Discovery Notice, Dkt. 248 @2,3.

¹⁵ *Id.* @ 6.

¹⁶ *Id.*

¹⁷ On May 4, 2016 the government provided additional discovery materials – Production 5_8 – that contain the documents related to SBP-1A-088, SBP-1A-089, SBP-1A-090, and SBP-1A091.

PROBLEMS CONDUCTING SEARCHES OF THE DISCOVERY

While the Summation Database was being loaded, the ability to search and canvass the material was an extremely tedious and time consuming process. For example, a “keyword search” often took as long as forty-five minutes before returning any results. On March 31, 2016 the lag time of the searches was reduced to less than a minute per search.¹⁸ As more of the productions were loaded, however, many of the previously performed searches had to be repeated to account for the new material. Further, because the defense didn’t have access to the government-assigned Bates labels, it was difficult to determine which files included in a specific search were newly added and which files were returned in previous searches.

Additionally, many of the files on Summation are not “searchable.” Most of the PDF and Word files are searchable but the text of many other types of files cannot be recognized by a keyword search. For example, those files that were uploaded as a tif, img, jpeg, png, avi, gif, bmp, or another image-type format are almost never included in the results of any “term search.” Currently, Summation contains approximately 3,482,204 files that are not “fully searchable.”¹⁹ As discussed above, the PD’s office is currently working on recoding of fields and reprocessing the load files. Once completed, many of the 3.5 million unsearchable files should be fully readable and searchable.

CONCLUSION

Counsel and the PD’s Office have worked tirelessly to address the issues pertaining to the discovery and remain committed to identifying and remedying problems as they arise and as counsel become aware of their existence. The issues detailed in this filing are being addressed but additional problems and issues may be identified as the parties continue to implement and test the proposed solutions and/or once all the productions have been fully uploaded. As such, it may be necessary for Counsel to supplement this report at a later date in order to best update the Court.

¹⁸ Summation still malfunctions on occasion due to the volume of the material contained on the database and these errors often cause delays in accessing the files.

¹⁹ 488,697 tif files; 1,392,988 msg; 461,962 unknown “?” files; 327,856 jpg files & 126,252 jpeg files; 110,429 gif files; 21,533 bmp files; 277,112 avi files.